DESCRIPTION

cDNA Fragment of Causative Gene of Spinocerebellar Ataxia
Type 2

TECHNICAL FIELD

The present invention relates to cDNA fragments of
the causative gene of spinocerebellar ataxia type 2
(hereinafter also referred to as "SCA2"), proteins
encoded thereby, antibodies corresponding to the proteins,
and antisense nucleic acids of the above-mentioned cDNA
fragments.

BACKGROUND ART

SCA2 is an autosomal dominant, neurodegenerative
disorder that affects the cerebellum and other areas of
the central nervous system.

It has recently been discovered that the causative
genes of 6 neurodegenerative diseases including
dentatorubral-pallidoluysian atrophy (DRPLA) have more
CAG repeats than the normal genes. That is, the numbers
of CAG repeats in the causative genes of the
neurodegenerative diseases are 37 to 100, while those in
the normal genes are less than 35.

It has been suggested that the causative gene of
SCA2 has an increased number of CAG repeats (Trottier, Y.
et al. *Nature*, 378, 403-406 (1995)). However, since the
causative gene of SCA2 has not been identified, and since
its nucleotide sequence has not been determined, SCA2
cannot be diagnosed by a genetic assay.

DISCLOSURE OF THE INVENTION

An object of the present invention is to provide a sequence-determined cDNA fragment of the causative gene of SCA2. Another object of the present invention is to

5   provide a protein produced by the causative gene of SCA2. Still another object of the present invention is to provide an antibody specific to the above-mentioned protein, which antibody is useful for therapy and diagnosis of SCA2. Still another object of the present

10   invention is to provide an antisense of the causative gene of SCA2, which is useful for therapy of SCA2.

The present inventors intensively studied to discover a *Tsp* E1 fragment with a size of 2.5 kb in which the number of CAG triplet is increased only in SCA2

15   patients, and partial sequence thereof was determined. Human cDNA library was screened using as probes the oligonucleotides that respectively hybridize with the regions between which the CAG triplet repeats are interposed, and a cDNA fragment which hybridizes with

20   both of these two probes was cloned. Using this cDNA fragment as a probe, human cDNA library was screened and a plurality of cDNA fragments which hybridize with the probe were cloned. Sequencing the cDNA fragments revealed that these cDNA fragments overlap with each

25   other. To sequence the 5'-end and 3'-end regions, RACE (rapid amplification of cDNA ends) was performed. Further, to sequence the 5'-end region, RT-PCR was

performed, thereby succeeding in sequencing the full length of the cDNA of the causative gene of SCA2.

That is, the present invention provides a nucleic acid fragment comprising a nucleic acid region encoding an amino acid sequence shown in SEQ ID NO: 1 (provided that the number of repeat units of Gln from the 166th to 188th amino acid varies between 15 and 100). The present invention also provides a protein having an amino acid sequence encoded by the nucleic acid fragment according to the present invention. The present invention further provides an antibody which undergoes antigen-antibody reaction with the above-mentioned protein. The present invention still further provides an antisense nucleic acid having a size of not less than 15 bp, which hybridizes with a mRNA transcribed from the nucleic acid fragment according to the present invention so as to inhibit translation thereof.

By the present invention, a sequence-determined cDNA fragment of the causative gene of SCA2 was provided. The protein encoded by the nucleic acid fragment according to the present invention may be used for therapy of SCA2 and may be used as an immunogen for preparing an antibody useful for therapy and diagnosis of SCA2. Further, since the nucleotide sequence of the causative gene of SCA2 was determined by the present invention, antisense to this gene may now be designed. Still further, by the present invention, a recombinant vector comprising the nucleic

acid fragment according to the present invention which is incorporated in an expression vector that can express a desired gene in human body, which recombinant vector can express the nucleic acid fragment in human body, as well

5   as a method for expressing the nucleic acid fragment according to the present invention comprising introducing the recombinant vector into human body. Thus, the present invention is thought to largely contribute to the therapy and diagnosis of SCA2.

BRIEF DESCRIPTION OF THE DRAWINGS

10   Fig. 1 shows the nucleotide sequence$_n$ of the cDNA (SEQ ID NO:1) fragment according to the present invention together with the amino acid sequence encoded thereby, which nucleotide sequence was determined in the Examples of the present

15   invention.

Fig. 2 shows the continuation of Fig. 1.

Fig. 3 shows the continuation of Fig. 2.

Fig. 4 shows the continuation of Fig. 3.

Fig. 5 is a pedigree chart of the SCA2 patients who

20   donated the genomic DNAs used in the Examples.

Fig. 6 shows the sizes, positions and restriction sites of the genomic DNA fragments Tsp1 and Tsp2, and SCA2 cDNA obtained in the Examples of the present invention, as well as the size and position of each of

25   the obtained cDNA fragment.

Fig. 7 shows the distribution of the numbers of the CAG repeat units in the normal and SCA2 genes, which were

measured by using the (CAG)$_{55}$ probe.

BEST MODE FOR CARRYING OUT THE INVENTION

As described above, the nucleic acid fragment according to the present invention comprises the nucleic acid region encoding the amino acid sequence shown in SEQ ID NO: 1 in the SEQUENCE LISTING, provided that the number of repeating units of Gln from the 166th to the 188th amino acid varies between 15 and 100. The number of this repeat units is 15 to 25 in normal individuals and 35 to 100 in SCA2 patients. As is well-known, due to degeneration, there are a plurality of codons encoding one amino acid, and any nucleic acids which encode the amino acid sequence shown in SEQ ID NO: 1 are included within the scope of the present invention. The nucleotide sequence actually determined in the Examples below is shown in SEQ ID NO:1 and in Figs. 1-4. The way how the nucleotide sequence was determined and the fact that the cDNA having this nucleotide sequence is the cDNA of the causative gene of SCA2 are detailed in the Examples below.

The nucleic acid fragment according to the present invention may be cloned by the method detailed in the Examples below. Further, since the nucleotide sequence of the nucleic acid fragment according to the present invention was determined by the present invention, the nucleic acid fragment may be cloned by utilizing amplification by PCR using human cDNA library as a

template, or by hybridization using the PCR product as a
probe. In cases where it is difficult to amplify the
nucleic acid fragment by a single PCR, the nucleic acid
fragment may be divided into a plurality of regions and
5   the PCR products may be ligated by a conventional method
so as to clone the nucleic acid fragment.

By incorporating the nucleic acid fragment according
to the present invention into the multicloning site of a
commercially available expression vector, and by
10 transforming host cells with the obtained recombinant
vector, the protein encoded by the nucleic acid fragment
may be produced. Host-vector systems for expressing an
arbitrary gene in a host cell is well-known in the art
and a number of host-vector systems are commercially
15 available. Those skilled in the art may easily express
the nucleic acid fragment according to the present
invention so as to produce the protein using such a
commercially available host-vector system. Such a well-
known method is described in, for example, D. M. Glover,
20 DNA Cloning Volume III a practical approach, 1987, IRL
Press.

For example, according to a conventional method, the
nucleic acid fragment according to the present invention
may be amplified by the long PCR (LA PCR) (Hiroyuki MUKAI,
25 PROTEIN, NUCLEIC ACID, ENZYME, Vol.41, No.5, 585-594,
1996) using primers in which restriction sites are
introduced, and the amplified product is digested with a

restriction enzyme. The digested amplified product is then inserted into the multicloning site of a commercially available plasmid vector pGEX (PHARMACIA) and ligated to obtain a recombinant vector, followed by

5 transforming *E. coli* DH5α (GIBCO BRL) by the conventional calcium chloride method. Transformants are then selected based on the drug resistance (ampicillin resistance) and the protein encoded by the causative gene of SCA2 may be recovered from the transformants. In cases where this

10 vector is used, the desired protein is obtained as a fusion protein with GST (Glutathion S-Transferase), so that the protein may easily be detected using a commercially available anti-GST antibody (PHARMACIA).

Since the active protein produced by the gene of

15 normal individuals (the number of CAG repeats is 15 to 25) is thought to have the normal function, SCA2 may be treated or alleviated by administering the normal protein to SCA2 patients. However, since SCA2 is autosomal dominant, it is necessary to simultaneously block the

20 causative gene of SCA2 of the patient with the antisense described below.

By immunizing an animal with the above-described protein or an antigenic fragment thereof, an antibody specific to the protein or the fragment thereof may be

25 recovered from the animal by a conventional method. The antibody may be a polyclonal antibody or a monoclonal antibody, and the monoclonal antibody may be prepared by

a well-known method.

It is theoretically difficult to construct an antibody using the polyglutamine chain encoded by the CAG repeat, which antibody has varying reactivities with the polyglutamine chain depending only on the lengths of the polyglutamine chain. However, by using the full length protein which was clarified by the present invention, an antibody which recognizes the stereoscopic difference between the patient's protein and the normal protein may be prepared. Further, by virtue of the present invention, since an antibody may be prepared by using a region of the SCA2 product, which is common to the normal individuals and the patients, a control is available in any assay system using an antibody specific to the protein according to the present invention. Using such an antibody, a simple assay such as plate method may be attained by a conventional method.

An affinity column may be prepared by immobilizing the antibody according to the present invention on agarose gel (e.g., Sephadex (trademark) of PHARMACIA), polystyrene beads or the like and by packing the resultant in a column. By passing a body fluid (blood, serum, spinal fluid or the like) through the affinity column, the protein produced by the causative gene of SCA2 of the patient may be obtained. By eluting the protein and by measuring the molecular weight thereof, diagnosis of SCA2 may be attained. This is because that

the protein of a SCA2 patient has a molecular weight

larger than that of the protein of a normal individual

because the number of CAG repeat units of the SCA2

causative gene in the patient is larger than that of the

5   gene in the normal individual.

Possible gene therapies include the method by which

production of the abnormal SCA2 gene product is stopped

and the method by which normal SCA2 gene product is

introduced. The former method includes blocking of

10   translation of mRNA using an antisense. The latter

method includes administration of the SCA2 gene product

prepared *in vitro*, and introduction of the nucleic acid

fragment into the cells. Since in patients suffering

from SCA2 caused by extension of the CAG repeat, the

15   abnormal protein is dominant, it is thought that the

desired effect may be obtained by simultaneously

performing both of the above-mentioned methods. In this

case, the antisense is designed such that it hybridizes

with a region not affecting the activity of the SCA2

20   product, and the region is removed from the normal SCA2

gene to be introduced, thereby assuring that the

antisense does not inhibit the production of the normal

protein. By designing the antisense and the normal SCA2

gene as mentioned above, the antisense and the SCA2 gene

25   may be incorporated in the same vector and the vector may

be introduced into cells.

The antisense nucleic acid according to the present

invention hybridizes with the mRNA transcribed from the nucleic acid fragment according to the present invention, and has a size of preferably not less than 15 bp and not more than the full length of the coding region of the

5 nucleic acid fragment, more preferably not less than 50 bp and not more than the full length of the coding region of the nucleic acid fragment. Although the antisense nucleic acid preferably has a nucleotide sequence completely complementary to the entire mRNA or a part

10 thereof, transcribed from the nucleic acid fragment according to the present invention, those having homologies to the degree that they hybridize with the mRNA *in vivo* are within the scope of the present invention. By administering the antisense according to

15 the present invention to an SCA2 patient, the causative gene of SCA2 may be blocked. By administering the above-mentioned protein from a healthy individual to the patient under this condition, SCA2 may be treated or alleviated. The dose of the antisense nucleic acid may

20 be appropriately selected depending on the conditions of the patient, and may usually be 0.001 mmol to 1000 mmol per day per 1 kg of bodyweight.

Known methods for introducing the gene into the body include methods in which Retrovirus or Adenovirus is used

25 as a vector (Yasuhiro SETOGUCHI, Experimental Medicine, Vol.12, No.15 (extra edition) 1994, pp.114-121; Hiromi KANEGAE et al., Experimental Medicine, Vol.12, No.15

(extra edition) 1994, pp.34-40), methods using liposomes, and methods using fusion liposomes (membrane-fused liposome and the like). Among these, since Adenovirus expresses the introduced gene without proliferation of

5   the target cells, the method utilizing Adenovirus is most appropriate as the method for introducing the nucleic acid into nervous system. More particularly, E3 and a part of E1a of the DNA of wild type Adenovirus type 5 are removed and the nucleic acid in double-stranded form is

10   introduced into the virus together with an expression unit such as a promoter. The recombinant virus vector is proliferated in 293 cells expressing E1a and E1b genes, originated from human fetal kidney to prepare a virus liquid, and the virus liquid is inoculated. The virus

15   genome incorporated into the nucleus of the cell exists outside the chromosomes without replication. Since the virus genome is not lost by cell division in nerve cells, expression of normal SCA2 gene may be maintained for 1 to 3 months.

20   The present invention will now be described more concretely by way of examples thereof. It should be noted that the present invention is not restricted to the following examples.

Example 1   Preparation of $(CAG)_{55}$ Probe

25   A genomic DNA segment of DRPLA gene containing a CAG repeat with 55 repeat units was amplified from the genomic DNA of a patient with DRPLA (Koide, R. et al.,

Nature Genet., 6, 9-13 (1994)) and was subcloned into a

plasmid vector, pT7Blue T(p-2093). The p-2093 plasmid

contains the (CAG)$_{55}$ (SEQ ID NO.6) and the flanking sequences. That is,

the plasmid contains the sequence of 5'-CAC CAC CAG CAA (SEQ ID NO.3)

CAG CAA (CAG)$_{55}$ CAT CAC GGA AAC TCT GGG CC-3'. Using a

pair of oligonucleotides 5'-CAC CAC CAG CAA CAG CAA CA- (SEQ ID NO.5)

3' (SEQ ID NO.4) and 5'-biotin-GGC CCA GAG TTT CCG TGA TG-3', PCR was

performed in a total volume of 16 µl containing 10 mM

Tris-HCl, pH8.3, 50 mM KCl, 1.5 mM MgCl$_2$, 2M N,N,N-

trimethylglycine, 0.1 mM TTP, 0.1 Mm dCTP, 0.1 mM dGTP,

9.25 MBq of [α-^{32}P]dATP (222 TBq/mmol), 0.5 µM each of

the two primers, 0.3 ng of plasmid DNA (p-2093) and 2.0 U

of Taq DNA polymerase (Takara Shuzo, Kyoto, Japan).

After an initial 2-min. denaturation at 94°C, PCR was

performed for 30 cycles consisting of 1-min. denaturation

at 94°C, 1-min. annealing at 54°C and 3-min. extension at

72°C, followed by a final extension at 72°C for 10 min.

A single-stranded (CAG)$_{55}$ probe was isolated using

streptavidin-coated magnetic beads (Dynabeads M-280,

Streptavidin;Dynal AS, Oslo, Norway) on which 20 µl of

streptavidin is coated. That is, after washing of the

PCR products immobilized on the magnetic beads with 40 µl

of a solution containing 5 mM Tris-HCl (pH 7.5), 0.5 mM

EDTA and 1 M NaCl, the non-biotinylated strand containing

the radio-label was separated from the biotinylated

strand by incubation in 50 µl of 0.1 M NaOH for 10 min.

The resultant supernatant was directly added to the

hybridization solution described below.

Incidentally, using the single-stranded $(CAG)_{55}$ probe prepared as described above, Southern blot analysis was carried out on the androgen receptor genes containing 9, 22, 43 and 51 CAG repeat units, respectively. As a result, the $(CAG)_{55}$ probe strongly hybridized with the genes having 43 and 51 CAG repeats units, respectively, but scarcely hybridized with the gene having 22 CAG repeat units, and did not hybridize at all with the gene having 9 CAG repeat units (K. Sanpei et al., Biochemical and Biophysical Research Communications, Vol.212, No.2, 1995, pp.341-346). Thus, by using this probe, hybridization may be selectively attained only with DNAs containing a number of (e.g., 35 or more) CAG repeat units if the hybridization conditions are appropriately selected.

(2) Determination of Nucleotide Sequence of SCA2 Gene

Fig. 5 shows a pedigree chart of SCA2 patients. In this pedigree chart, males are represented by squares and females are represented by circles. SCA2 patients are represented by black squares or circles, and unaffected persons are represented by white squares or circles.

High-molecular-weight genomic DNA (15 μg) was digested with 100 U of *TspEI* (Toyobo, Osaka, Japan), electrophoresed through 0.8% agarose gels and transferred to nitrocellulose membranes. The membranes were hybridized with the $(CAG)_{55}$ probe described above.

Hybridization was performed in a solution containing 2.75 x SSPE (1 x SSPE=150 mM NaCl, 10 mM NaH_2PO_4, 1 mM EDTA), 50% formamide, 5 x Denhardt's solution, 100 ng/ml of sheared salmon sperm DNA and the $(CAG)_{55}$ probe (6×10^6 cpm/ml) at 62°C for 18 hours. After the hybridization, the membranes were washed with 1 x SSC (150 mM NaCl, 15 mM sodium citrate) containing 0.5% SDS at 65°C for 0.5 hours. The membranes were autoradiographed for 16 hours to Kodak Bio Max MS film at -70°C using an MS intensifying screen.

As a result, 2.5 kbp *TspEI* fragment hybridized with the probe was detected only in all of the SCA2 patients.

Genomic DNA (270 µg) from an SCA2 patient (individual 7 in Fig. 5) was digested by *TspEI* and subjected to agarose gel electrophoresis. Genomic DNA fragments including the 2.5 kb *TspEI* fragment were cloned into an *EcoRI*-cleaved λZAPII vector. The genomic library was screened using the $(CAG)_{55}$ probe under the hybridization condition described above. A genomic clone, *Tsp-1*, containing an expanded CAG repeat was isolated.

After removal of the probe, the above-described genomic library was screened again using the *Tsp-1* as a probe, which was labeled by the random priming. Hybridization was carried out in a solution containing 5 x SSC, 1 x Denhardt's solution, 10% dextran sulfate, 20 mM sodium phosphate, 400 µg/ml human placental DNA and the *Tsp-1* probe at 42°C for 18 hours. After the

hybridization, the membranes were washed finally in 0.1 x

SSC - 0.1% SDS at 52°C for 0.5 hours. The membranes were

autoradiographed for 24 hours to Kodak Bio Max MS films

at -70°C using an MS intensifying screen. As a result, a

5   genomic clone, *Tsp-2*, originated from a normal allele was

isolated.

The SmaI-ApaI fragment (630 bp) of Tsp2 was

sequenced and oligonucleotides F-1 (5'-CCC TCA CCA TGT

CGC TGA AGC-3') (SEQ ID NO: 7) and R-1 (5'-CGA CGC TAG AAG GCC GCT G-

10   3') (SEQ ID NO: 9) were designed such that the CAG repeat units are

sandwiched between the oligonucleotides (see Fig. 1).

Using oligonucleotides F-1 and R-1 as probes, human

procephalic cortex cDNA library (STRATAGENE) was screened.

Hybridization was performed in a solution containing 6 x

15   SSC, 10 x Denhardt's solution, 0.5% SDS, 0.05% sodium

pyrophosphate, 100 ng/ml of sheared salmon sperm DNA and

end-labeled oligonucleotide probes at 55°C for 18 hours.

After the hybridization, the membrane was finally washed

with 6 x SSC containing 0.5% SDS and 0.05% sodium

20   pyrophosphate at 55°C for 0.5 hours. A cDNA clone Fc1

with a size of 4.0 kb which hybridized with the both

probes was obtained. The nucleotide sequences of Fc1,

Tsp1 and Tsp2 were determined and compared. As a result,

the nucleotide sequences in the vicinities of the CAG

25   repeat units were identical except for the number of the

CAG repeat units. Restriction maps of Tsp1 and Tsp2, as

well as the sizes and positions of Fc1 and other

fragments hereinbelow described, are shown in Fig. 6.
Using Fc1 or a fragment isolated by the screening later
described as a probe, human cDNA libraries (human
procephalic cortex, human fetal brain, human brain and

5   brain stem) were screened to isolate cDNA clones Fc2,
Fb14, B4, C6 and C19 (see Fig. 6). To identify the 5'-
end of Fc1, 5'-RACE (Frohman, M.A. et al, Proc. Natl.
Acad. Sci. USA 85, 8998-9002 (1988)) was performed using
5'-RACE-Ready cDNA (Clonetech, Palo Alto, CA, USA).

10   Primer R-1 was used for the first PCR, and Primer R-2
(5'-CTT GCG GAC ATT GGC AGC C-3', see Fig. 1) was used _(SEQ DNO.9)_
for the second PCR. In both PCRs, F-1 (see Fig. 1) was
used as the forward primer. The 5'-RACE product (5R1)
having the size of 350 bp was subcloned into pT7Blue T

15   vector (pT7Blue T-vector (5R1)). The identification of
5R1 was confirmed by the overlapping with the nucleotide
sequences of Fc1, Tsp1 and Tsp2. To identify the 3'-end
of the cDNA, 3'-RACE was performed using 1 μg of
poly(A)$^+$mRNA extracted from human brain as a template and _(SEQ D NO.11)_

20   Primer F-13 (5'-TTC TCT CAG CCA AAG CCT TCT ACT ACC-3',
see Fig. 3) as a primer. The obtained 3'-RACE product
(3R1) with a size of 1300 bp was subcloned into pT7Blue T
vector (pT7Blue T-vector (3R1)).

     To investigate the 5'-end region of the cDNA,

25   reverse transcription PCR (RT-PCR) was performed. That
is, total RNAs extracted from an autopsy from human brain
were digested by RNase-free DNase (PROMEGA) (Onodera, O.

et al., Am. J. Hum. Geent. 57, 1050-1060(1995)). As the

primers for the PCR, F1006 (5'-TAT CCG CAG CTC CGC TCC

(See ID NO:11)

C-3', see Fig. 1) and R1002 (5'-AGC CGG GCC GAA ACG CGC

(See ID NO:12)

CG-3') were used. PCR was performed in a solution with a

total volume of 20 μM, which contained 5 pmol each of the

each primer, 10 mM Tris HCl (pH8.3), 50 mM KCl, 1.5 mM

$MgCl_2$, 1.7M N,N,N-trimethylglycine, 200 μM each of dATP,

dCTP and TTP, 100 μM of dGTP, 100 μM of 7-deaza dGTP and

2.5 U of Taq polymerase (TAKARA SHUZO). After carrying

out the initial denaturation at 96°C for 2 minutes, a

cycle of a denaturation step at 96°C for 1 minute, an

annealing step at 65°C for 1 minute and an extension step

at 72°C for 1 minute were repeated 30 times, and a final

extension step at 72°C for 5 minutes was performed,

thereby carrying out the PCR. As a result, a clone 5R1

which extends upstream of 5R1 by 246 bp was obtained (see

Fig. 6).

In Fig. 6, the hollow regions in the Tsp1 and Tsp2

fragments indicate the regions which exist in SCA2 cDNAs.

The hollow regions in the SCA2 cDNA shows coding regions.

The CAG repeating regions are shown as solid boxes.

Restriction sites TspE1 (T) ، NotI (N)، Sac II (S)، Sau3AI

(Sa) ، Sma I (Sm)، Eco52I (E52)، Apa I (Ap)، AccI (Ac) ،

BamHI (B) ، XhoI (X)، EcoRI (E) and Pst I (P) are shown.

The size and position of each cDNA clone are shown below

the consensus SCA2 cDNA.

In this example, nucleotide sequences of double-

stranded DNAs were determined by the dideoxynucleotide
chain termination method (Sanger, F. et al. Proc. Natl.
Acad. Sci. USA 74, 5463-5467(1977); Chen E.Y. et al, DNA
4, 165-170 (1985)) using a double-stranded plasmid DNA as

5   a template. To determine the nucleotide sequences of the
CAG repeating regions and their flanking regions, genomic
fragments containing the CAG repeating regions were
amplified by PCR using biotinylated F-1 and RS-1 (5'-CCT
(SEQ ID NO: 13)
CGG TGT CGC GGC GAC TTC C-3'). PCR was performed in a

10 solution with a total volume of 25 µl, which contained
0.25 µM each of the each primer, 10 mM Tris HCl (pH8.3),
50 mM KCl, 2.0 mM $MgCl_2$, 1.7M N,N,N-trimethylglycine, 200
µM each of dNTP, 200 ng of the genomic DNA and 1.25 U of
Taq polymerase (TAKARA SHUZO). After carrying out

15 initial denaturation at 95°C for 1 minute, a cycle of a
denaturation step at 95°C for 2 minutes, an annealing
step at 62°C for 1 minute and an extension step at 72°C
for 1 minute was repeated 32 times, and a final extension
step at 72°C for 5 minutes was performed, thereby

20 carrying out the PCR. Biotinylated chains were recovered
using streptavidin-coated magnetic beads and were
directly sequenced.

Based on the nucleotide sequences of the above-
mentioned cDNA clones, a consensus SCA2 cDNA sequence

25 with a length of 4351 bp excluding the poly A tail was
determined (SEQ ID NO:1, Figs. 1-4, see Fig. 6). In SEQ
ID NO: 1, the region from 4352nt to 4367nt is the poly A

tail, and the number of "A" is not restricted to that shown in SEQ ID NO: 1. It was confirmed that the poly A tail exists at the same location in C19, B4 and 3R1 which were independent cDNA clones.

5   Example 2   Measurement of CAG Repeat Units in Sample

Numbers of CAG repeat units were determined by polyacrylamide gel electrophoresis analysis of PCR products obtained using the primer pair of F-1 and R-1. PCR was performed in a total volume of 10 µl containing 10   10 mM Tris-HCl, pH 8.3, 50 mM KCl, 2.0 mM $MgCl_2$, 1.7 M N,N,N-trimethylglycine, 111KBq of $[\alpha-^{32}P]$dCTP (111 Tbq/mmol), 30 µM dCTP, and 200 µM each of dATP, dGTP and TTP, 0.25 µM each of the two primers, 200 ng of genomic DNA and 1.25 U of Taq DNA polymerase. After an initial 15   2-min denaturation at 95°C, PCR was performed for 32 cycles of 1-min denaturation at 95°C, 1-min annealing at 60°C and 1-min extension at 72°C, followed by a final extension at 72°C for 5 min. Sequence ladders obtained using the cloned genomic segments of the SCA2 gene, which 20   contain various sizes of CAG repeats, were used as size markers. For normal alleles containing one or two CAA interruptions, the numbers of the CAA units were included in the CAG repeat size. For SCA2 alleles having expanded CAG region, the above-mentioned insert sequence 25   immediately after the CAG region was not included in the size of the CAG region.

By the above-described method, the numbers of the

CAG repeat units of normal individuals (286 chromosomes) and 10 pedigrees of SCA2 patients (34 SCA2 chromosomes) were determined. The results are shown in Fig. 7. In Fig. 7, open bars indicate the results of the normal genes and solid bars indicate the results of the SCA2 genes.

As is apparent from Fig. 7, in all of the normal genes, the numbers of the CAG repeat units were not more than 24, while in all of the SCA2 genes, they were not less than 35. Thus, it was confirmed that the cDNA identified as described above is the cDNA of the causative gene of SCA2.